Article:
Olson, E.T. (1997) The Ontological Basis of Strong Artificial Life. Artificial Life, 3 (1). pp. 29-39. ISSN 1064-5462

A slightly different version of this paper appeared in Artificial Life 3 (1997): 29-39.

Computer-Generated Life

Eric T. Olson

To create a little flower is the labour of ages.

Blake

i.

Dr. Frankenstein, in Mary Shelley's novel, created a living human being in his laboratory. Impressive though this accomplishment may have been, though, Frankenstein's monster was made from materials that were already living, or nearly so, before they came into his hands. His "creation" of life was little more than an elaborate bit of surgery and resuscitation. It would be of greater philosophical interest if someone could create a living organism out of non-biological materials--out of simple organic molecules of the sort that were present on the young earth, for example, or out of wholly inorganic chemicals, or even out of nuts and bolts and wires. This would truly be a case of artificial life.

However, an increasing number of computer scientists claim that they can do something even more surprising than this. They say that it is possible to create genuine living organisms without getting one's hands dirty, simply by programming a computer in the right way.

Here are some examples of this astonishing claim. C.G. Langton, one of the founders of the new science of artificial life, writes,

> The ultimate goal of the study of artificial life would be to create 'life' in some other medium, ideally a virtual medium where the essence of life has been abstracted from the details of its implementation in any particular hardware. We would like to build [computer] models that are so lifelike that they would cease to be models of life and become examples of life themselves (1986, 147).

Langton clearly believes that this can be done. "Any definition or list of criteria broad enough to include all known biological life," he argues, "will also include certain classes of computer processes which, therefore, will have to be considered to be 'actually' alive" (1992b, 19).

Claus Emmeche describes Thomas Ray's (1992) programming work the work by saying,

> with the help of computers he could piece together fragments of computer programs (i.e., instructions) and turn them into artificial organisms that did not just resemble life, but that theoretically speaking were just as alive as real animals and plants." (1994b, 3) [1]

Others who have expressed at least guarded support for Computer-Generated Life include Farmer and Belin (1992), Keeley (1994), Levy (1992), Pattee (1989), Rasmussen (1992), Spafford (1992), and Taylor (1992).

I am going to call the claim that one can create a living organism simply by programming a computer in the right way "Computer-Generated Life". It is important to distinguish this from two less radical positions that share the name "Artificial Life". First there is the perfectly respectable claim that computer models of complex biological systems can provide biologists with novel ways of learning about the living world. Computerized simulations of life may also enable us to learn about possible forms of life that are not in fact known to exist, including organisms based on structures other than carbon chemistry. These claims about the modelling and simulating powers of artificial-life programs often go by the name of "Weak A-life" (e.g. Langton 1992b). "Strong A-life", on the other hand, is the bold hypothesis that it is possible to go beyond simulation and literally create living organisms with the help of computers. These may be chemical creatures in a test tube, or they may be mechanical automata of some sort. "Computer-Generated Life" goes beyond Strong A-life by saying that one can create a living thing without any materials at all, just by manipulating information.

This claim raises a number of philosophical questions. [2] Could today's or tomorrow's artificial-life programmers really create something living? What sort of thing can they create at all? What is a living organism, anyway? Is life essentially a computational phenomenon? And if so, would that mean that living organisms were somehow made out of information rather than out of matter? What is the relation between natural organisms and computer-generated ones? When is a computer model of some phenomenon an instance of that phenomenon itself? What sort of environment would these computer-generated organisms inhabit, and what is the relation between that environment and our own? What ontological category would computer-generated organisms belong to? Are they supposed to be material objects? Events or processes? Platonic complexes of pure information? Or are the traditional ontological categories of the philosophers adequate to account for this new phenomenon?

I shall try to address only a few of these questions. My main purpose will be to show that the hypothesis of Computer-Generated Life rests on shaky conceptual foundations. For there to be life is for there to be living organisms. (Life is a property, and the things that bear that property are living things: organisms.) I believe that the advocates of Computer-Generated Life have not thought carefully about just what sort of things these computer-generated organisms are supposed to be. It is not clear what things they claim to be able to create, and which they take to be alive.

They may point to a pulsating image on the monitor and say, "That is alive." But what thing are they pointing at? Is it a material object like a dog or a cat, that is literally, physically inside the computer? A physical process going on there that is "bioelectrical" rather than biochemical? An abstract complex or pattern of information, such as a program? Or is there some other category of thing that computer-generated organisms could belong to? Friends of Computer-Generated Life argue that the things they create, or could create, are genuinely alive; but we cannot evaluate this claim until we know what they have created.

We are often told that the would-be organisms in question are "certain classes of computer processes" (Langton 1992b), or exist "in silico" as opposed to "in vitro" (Langton 1989b, Emmeche 1994b, etc.), or are "constructed entirely out of machine instructions" (Ray 1992, 374), or that their "essence is information" (Levy 1992, 6). But this is not very helpful, for it is not clear what sort of thing a class of computer processes or a machine instruction is supposed to be; and no such explanation is offered.

The reader may suspect that these questions are tainted with philosophical dogma (though they are not entirely without precedent; see Bedau 1992, 496). Why should computer-generated organisms fit neatly into any of those ontological categories--material object, event or process, or abstract object? But the same point can be made by asking plainer questions that do not involve those categories: Are these organisms made of matter? If so, what matter; and if not, what are they made of? Are they literally inside the computer? If not, where are they? Or are they located in space and time at all? What is their relation to the hardware of the computer running the artificial-life program? What prevents a computer-generated organism from biting me? If there really are computer-generated organisms, these questions ought to have answers--particularly if those organisms are supposed to be comparable with natural living things like elephants and giraffes.

I shall discuss several possible answers to these questions--all the ones that I can think of--and try to show that they are all problematic. I do not claim to show that Computer-Generated Life is absurd. But its defenders need to think more about what they are saying.

ii.

Let us consider first the picture of computer-generated organisms that I have heard most often from their defenders: that they are abstract complexes or patterns of pure information, as opposed to concrete, physical objects or processes that exist or occur at particular places and times. Here is Langton once again:

> Computers themselves will not be alive, rather they will support informational universes within which dynamic populations of informational "molecules" engage in informational "biochemistry." (1989b, 39; see also 1986)

These purely informational organisms would be something like computer programs--though they need not be the programs themselves, since a single program might generate more than one organism, or an organism and a purely informational "environment" for it; and there might be portions of an artificial-life program that have nothing to do with the organism it produces.

A computer program is abstract in that it is distinct from any of its particular individual instances or copies stored on disks, written out on paper, or running on a piece of hardware. When you copyright a program, you don't copyright any particular copy of the program; you copyright the information of which those copies are instances. It is a mistake to ask for a computer program's location or shape or chemical composition, for it is not a physical thing at all, and it is located in space and time only insofar as it has individual, concrete instances at particular times and places. It is a universal. Computer-generated life would represent a purely Platonic or Pythagorean biology.

Could such a thing be a living organism? It is hard to see how it could be. Pure information is not capable of change. It doesn't do anything. Only the particular, concrete items that realize or instantiate a piece of information can change. Of course we may speak of corrupting a program, or a similar piece of information, by changing some of its coded instructions; but that affects only a particular copy of it. The program itself--the thing that is copyrighted--is unchanged, and continues to run properly in its uncorrupted instances.

On the other hand, change is part of the very essence of life. A living organism comes into being, grows and develops, metabolizes, reproduces (perhaps), and dies, leaving behind something like inert matter. That may be something that a particular instance of a complex of pure information can do, but not the information-complex itself. Saying that a computer program is alive is like saying that the species Elephas maximus is a living thing--rather than any of the individual members of that species tromping around India.

If this argument fails to convince, there is another. It is not only certain electronic computers that are associated with lifelike complexes of information. The same is true of natural organisms. A computer running an artificial-life program is surely no more lifelike than an elephant, which is a paradigm case of a living thing. If computer programmers can produce purely informational life, then there is also naturally occurring purely informational life. An elephant realizes or conforms to many abstract patterns of information; and whatever it is about a computer program that makes it a living thing would surely make some of those elephantine patterns living things as well. (Perhaps some of them could even be translated into some programming language and run on a digital computer.) So if there are purely informational artificial organisms, there are also purely informational elephants.

This should come as no surprise to the defenders of Computer-Generated Life, if they see their creations as abstract patterns of information. But it has a surprising consequence for their view. The elephant that eats 300 pounds of vegetable matter and produces twelve gallons of urine each day is not a pattern of pure information. It is a dangerous, five-ton material object. Granted, that five-ton heap of matter may be a living thing rather than a non-living thing because it instantiates a special kind of information-pattern. At any rate the difference between living material things and nonliving material things is a matter of the way their component parts interact with one another and with their environment, and not a matter of what sort of stuff they are made out of. Information is part of the elephant's essence, so to speak. But we should not conclude from this (as some seem to do) that elephants are patterns of pure information and not material objects, or that it is the pattern and not the material thing that is alive.

So for every purely informational elephant, there is a physical, flesh-and-blood elephant that we can see and feel and smell. In that case, there ought to be some concrete, physical living thing corresponding to each purely informational organism produced by a computer as well. We should expect there to be a physical organism inside the computer that instantiates the artificial-life program in the way that the physical elephant instantiates its own lifelike information-complex. So if you can produce a purely informational organism by programming a computer, you can also produce a concrete, physical organism in that way. Computer-generated life is possible only if it is possible to produce concrete, physical living organisms by programming a computer--even though there may also be purely informational organisms. So let us turn our attention to this second picture of computer-generated life.

iii.

Several writers seem to think of computer-generated organisms as concrete, physical things that are literally inside the computer, like a hamster in a cage or a flower in a pot. Brian Keeley, for example, writes,

> [I]t is in virtue of blip world's physical properties (not its computational properties) that it exhibits relevantly biological behavior. While it is true that the medium in which this behavior is found is a "computer," we should never forget that our computer is not some kind of Platonic "purely computational system"; it is a very down-to-earth physical system, a machine. (1994, p. 578; see also Emmeche 1994b, x, 18)

(Blip world is the product of an artificial-life program modeled on Tom Ray's Tierra [1992].) And Robert Davidge (1992) says that computer-generated organisms are made up of transistors in much the same way as natural organisms are made up of molecules.

This picture would enable one to avoid the criticism brought by some writers that putative computer-generated organisms are "just ungrounded symbol systems that are systematically interpretable as if they were alive; in reality they are no more alive than a virtual furnace is hot" (Harnad 1994, 539; see also Sober 1992). And as we saw earlier, it is hard to see how anything other than a physical object or event could have the sorts of causal properties, such as metabolism and interactions with their surroundings, that are essential to life.

So let us consider the view that computer-generated organisms are material objects, just as elephants are: things made of matter and having a mass, microphysical composition, size, and location. It may sound fantastic to suppose that programming a computer could be a way of bringing a material object into existence, or of turning a previously nonliving material object into a living one. If computer-generated organisms were physical objects, they would have to made out of materials found in the hardware of a computer. Presumably they would be spidery things made of bits of metal and silicon and the like. You could actually see them by looking into the computer's innards, and you could calculate their size and mass. Since they could not literally move except by expanding or contracting to assimilate more or less of the computer's hardware, they would be more vegetable than animal. When the program whose complex electronic structure brought those materials to life stopped running, the organism would literally die, leaving behind a wiry corpse.

Strange as this picture may appear, I believe that it is the most defensible way of understanding Computer-Generated Life. Atoms that are individually lifeless may make up a natural living organism, such as a bacterium, by being caught up in a special kind of chemical event. That chemical metabolism unifies those atoms into a living thing. In the same way, perhaps an electronic circuit with a structure and complexity similar to that of biochemical metabolism could unify some of the metal and silicon atoms found inside a computer into a living thing. The advocates of Computer-Generated Life would have discovered a biological distinction even more fundamental than that between prokaryotes and eukaryotes, namely that between chemical life and electronic life. A bold hypothesis--but not obviously absurd.

While it is bold to suggest that one could produce a living material object by programming a computer, no one would dispute that it is possible to produce a physical <u>event</u> or process in that way (except for those austere metaphysicians who deny that there are any events at all); and one might try to argue that computer-generated organisms are electronic events rather than material objects. Although they are concrete, particular things, events are not material objects. They are things that <u>happen to</u> material objects, or activities that material objects participate in. Of course, not just any event going on inside a computer is a plausible candidate for being a living organism. None of the events that are now taking place within my humble Macintosh Plus is a living thing. Living computer processes would have to have a kind and degree of organized complexity that ordinary, non-living ones lack, just as a real horse has a kind of organized complexity not present in a bronze statue of a horse. For a computer process to count as a living organism, you might think, it need only have a kind and degree of complexity comparable to that of at least the simplest possible natural organism.

For all I know there are or soon will be computer processes that meet this standard of organized complexity. However that may be, though, there are many <u>natural</u> processes that already measure up. If any electronic process is a living organism, the vast biochemical process that is the metabolism or individual biological life of an elephant is certainly also lifelike enough to be a living organism. [3] The elephant itself, though, is no more an event or process than it is a complex of pure information. The elephant's metabolism is something that happens to the elephant, an activity that the elephant's atoms are caught up in. Although it may be roughly the same size and shape as an elephant, the elephant's biological life does not weigh five tons, or have floppy ears and fist-sized teeth as material parts. The elephant and its metabolism are as different as the hardware of a computer and the electronic processes going on there.

But if both the elephant and its life or metabolism are alive in the very same sense, then there are two elephants whenever we thought there was one: the five-ton, material elephant, and the "process" elephant, consisting of the metabolic processes of the material elephant. Computer-Generated Life would have radical consequences for the biology of <u>natural</u> organisms. So the claim that computer-generated organisms are concrete, electronic processes seems even more difficult to believe than the view that they are physical objects made of metal and silicon.

At any rate, the "process" picture of Computer-Generated Life has no theoretical advantage over the "material-object" picture. The existence of living biochemical process in nature entails the existence of living material objects: you can't have an elephant metabolism unless you have a five-ton, material elephant. In the same way, we should expect each living computer process to correspond to a living material object, made from some of the wire and silicon inside the computer. So if there are living computer processes, it seems that there must also be living pieces of computer hardware.

iv.

I have argued that computer-generated organisms must be either abstract complexes of information not located in space and time, electrical events or processes, or physical objects made of metal and silicon. And I have argued that the first two options each entail the third. You may think that these categories are not exhaustive, and that we have been looking for artificial life in the wrong place. Let us consider one more alternative.

Some computer scientists seem to think that while computer-generated organisms are rather like physical objects, they are not literally physical things in the way that elephants are. They inhabit a different world from ours: a "symbolic world" or "virtual medium" of their own. They aren't made of quarks and electrons, but of completely alien materials to which mass, momentum, electric charge, and the other basic properties of "our" physics may not apply. Their world has its own laws of nature. Those laws need only be enough like ours to allow for the possibility of life (perhaps some analogue of the second law of thermodynamics must hold there, for example). Programmers create artificial organisms by creating an entire artificial world. That world is not part of the physical space inside the computer. It is an environment that is not spatially related to us at all. The images we see on the monitor are pictures of things going on in a different universe.

This may be what H.H. Pattee has in mind when he writes,

> [S]trong AL...can treat the symbolic domain of the computer as an artificial environment in which symbolic phenotypic properties of artificial life are realized. One might object that [this] is not a realization of life since environment is only simulated. But we do not restrict life forms to the Earth environment or to carbon environments. Why should we restrict it to non-symbolic environments? (1989, 65).

While Pattee does not explicitly endorse this view, he considers it a live option. (However, it may be that Pattee's "symbolic" environment is meant to be just a version of the "Platonic" picture of computer-generated organisms as abstract complexes of information, discussed earlier.) Steen Rasmussen, however, clearly thinks we can create artifical worlds as concrete and tangible as our own, with their own laws of nature:

> An artificial organism must perceive a reality R_2, which, for it, is just as real as our "real" reality, R_1, is for us....Assuming R_2 exists in a computer, its properties may be very different from the properties of R_1. From a logical point of view it is possible to create interactions in R_2 which do not in any direct way obey the physics in R_1. We can, thereby, create a more general physics in our universal machines than the physics we know (1992, 769f.) [4]

Now why should anyone think that she has created something beyond what is going on inside her computer when she writes and runs an artificial-life program? It is instructive to compare these programs with works of fiction, such as novels or animated films. Somebody might think that when Tolkien wrote <u>The Hobbit</u> he created a planet--made of rock, like our own--peopled with elves and goblins and other fantasic creatures; and that his novel is not just a story, but a true chronicle of actual events that took place there. That world is obviously not part of our own universe (certainly the laws of nature that govern our universe don't apply there). Nevertheless, one might say, it exists as a concrete thing just as our own world does. It is not enough for there to be a universe that would have existed whether Tolkien had written about it or not. He <u>created</u> that world, just as computer scientists create organisms when they program a computer in the right way.

This is a perfectly absurd way to understand what happens when someone produces a work of fiction. Not only would it make storytellers into gods with supernatural powers (and make them morally accountable for the sufferings of their creatures); it also faces technical difficultes, such as the fact that no work of fiction is specific enough to correspond to any one way a universe could be, and that some novels contain inconsistencies, and so could not be true histories of anything.

But the comparison between artificial-life programs and works of fiction raises a hard question: If artificial-life programmers create "virtual worlds" that their programs represent, must we also say that novelists or the makers of animated films create fictional worlds that their stories represent? (Remember that these "informational worlds" are not physically inside the computer; that is an option we discussed earlier.) Can we believe, on principled grounds, that there are concrete, flesh-and-blood worlds generated by computer programs, without also believing that there are concrete, flesh-and-blood fictional worlds? If not, we must either reject this picture of Computer-Generated Life or accept the "truth-in-fiction" claim.

Perhaps this analogy is misleading. One important difference between artificial-life programs and novels is that the programs give a far more detailed description than novels do. Typically there is just no saying what even the most precisely drawn fictional character would do in any situation not explicitly included in the story. Doyle spends some 1200 pages telling us about Sherlock Holmes; yet he did not provide us with any way of finding out what Holmes might have done with the Rodney King case. (Nor would it make any difference if those stories were incorporated into a computer program, unless many facts about Holmes were added that Doyle did not specify.) An artificial-life program, however, is founded on explicit axioms, which (ideally at least) generate a specific result for any possible specification of "initial conditions". These programs generate a wealth of detailed information about whatever it is that they are about. That is what makes artificial-life programs and other computer models so useful and interesting: if the information written into them corresponds accurately to real conditions, they can generate highly detailed and instructive predictions about real events. A computer model of a storm is likely to tell us far more about the storm, or about storms in general, than a novelist's description of a storm.

But this difference alone could not make it the case that artificial-life programs correspond to real events while fictional stories do not. Making a description more detailed and specific does not make it any more likely to be true of anything. Quite the opposite: a specific description is <u>less</u> likely to be true than an unspecific one. Anyone who thinks that artificial-life programs are relevantly different from works of fiction needs to provide us with some other reason for thinking so.

I may have been uncharitable in dismissing the "truth-in-fiction" hypothesis as absurd. Perhaps there really are hobbits and trolls and talking trees somewhere, which are as concrete as lions and tigers; and Tolkien created them when he wrote his novels. But in that case what is remarkable about the science of artificial life is not that it is capable of creating living organisms, for any storyteller can do that. What is remarkable is that the organisms those programmers create are more detailed or interesting or instructive than the creatures of fantastic literature. The value of this new science, then, lies in its capacity to simulate or model living systems, not in its capacity to create them. [5]

Notes

1. See Emmeche 1994a for his own defense of Computer-Generated Life.

2. These and others are surveyed by Mark Bedau (1992).

3. For a discussion of an organism's individual life or metabolism as a kind of activity, see van Inwagen 1990, sect. 9, and Young 1971, ch. 6.

4. See also Davidge 1992, 451. Poston and Fairchild (1993) claim that some virtual-reality applications work best when they are modeled on Aristotelian rather than Newtonian physics.

5. For valuable conversations on this topic I am indebted to Carlo Maley, John W. Manly, Kelly Salsbery, and several members of the Artificial Life Reading Group at MIT.

References

Bedau, Mark A. 1992 "Philosophical Aspects of Artificial Life", in Varela and Bourgine (1992), 494-503

Davidge, Robert 1992 "Looking at Life", in Varela and Bourgine (1992), 448-455

Emmeche, Claus 1994a "Is Life as a Multiverse Phenomenon?" in Langton (1994), 553-568

Emmeche, Claus 1994b The Garden in the Machine: The Emerging Science of Artificial Life, tr. by Steven Sampson, Princeton: Princeton University Press, 1994

Farmer, J. Doyne and Belin, Alletta 1992 "Artificial Life: The Coming Evolution", in Langton (1992a), 815-839

Harnad, Stevan 1994 "Artificial Life: Synthetic vs. Virtual", in Langton (1994), 539-552

Keeley, Brian 1994 "Against the Global Replacement: On the Application of the Philosophy of Artificial Intelligence to Artificial Life", in Langton (1994), 569-587

Langton, C.G. 1986 "Studying Artificial Life with Cellular Automata, Physica 22D, 120-149
___ ed. 1989a Artificial Life: the proceedings of an interdisciplinary workshop on the synthesis and simulation of living systems (Redwood city, CA: Addison-Wesley)
___ 1989b "Artificial Life", in Langton (1989a), 1-48
___ et al., eds. 1992a Artificial Life II (Reading, MA: Addison-Wesley)
___ 1992b "Introduction", in Langton (1992a), 3-24
___ ed. 1994 Artificial Life III (Reading, MA: Addison-Wesley)
Levy, Steven 1992 Artificial Life (New York: Vintage Books)
Pattee, H.H. 1989 "Simulations, Realizations, and Theories of Life", in Langton (1989a), 63-78
Poston, Timothy and Fairchild, Kim Michael 1993 "Virtual Aristotelian Physics", in IEEE Virtual Reality Annual International Symposium (Piscataway, NJ, IEEE), 512-518
Rasmussen, Steen 1992 "Aspects of Information, Life, Reality, and Physics", in Langton (1992a), 767-771
Ray, Thomas S. 1992 "An Approach to the Synthesis of Life", in Langton (1992a)
Sober, Elliott 1992 "Learning from Functionalism--Prospects for Strong artificial Life", in Langton (1992a), 749-765
Spafford, Eugene H. 1992 "Computer Viruses--A Form of Artificial Life?" in Langton (1992a), 727-745
Taylor, Charles E. 1992 "'Fleshing Out' Artificial Life II", in Langton (1992a), 25-38
van Inwagen, Peter 1990 Material Beings (Ithaca: Cornell University Press)
Varela, Francisco and Bourgine, Paul, eds. 1992 Toward a Practice of Autonomous Systems: proceedings of the first European conference on artificial life (Cambridge, MA: MIT Press, 1992)
Young, J.Z. 1971 An Introduction to the Study of Man (New York: Oxford University Press)